And

Reserving inter-domain lambda and compute resources across US and Japan

Gigi Karmous-Edwards
(MCNC)

Tomohiro Kudoh
(AIST)

ECOC 2007 workshop
Network for IT
September 16th, 2007

Outline

- G-lambda Overview
- Enlightened Computing Overview
- NLR Partnership
- Global Collaboration (G-lambda and Enlightened)
- 3-way Collaboration with PHOSPHORUS
- GLIF Control plane wg Update

G-lambda

Network as a resource

- In computing systems, hardware components such as CPUs, memories and storages are "resources".

- Resources are managed by resource managers (c.f. queuing system of computing center), and a user can use them while they are allocated to the user.

- To provide a stable service using "resources" connected to wide area network, network itself should be also considered as a "resource" and bandwidth should be allocated to users appropriately.

- To use network as a resource, an I/F to request bandwidth from users is required.
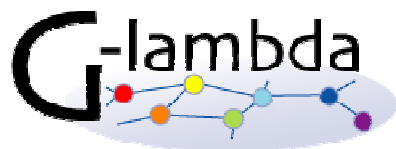
G- lambda project overview

- Joint project of KDDI R&D labs., NTT, NICT and AIST.
- G-lambda project has been started in December 2004.
- The goal of this project is to define a standard web services interface (GNS-WSI) between Grid resource manager and network resource manager provided by network operators.

National Institute of
Advanced Industrial Science
and Technology
AIST

NTT

NiCT

National Institute of
Information and
Communications
Technology

JGNⅡ

KDDI

KDDI R&D LABS

The G-lambda Team

AIST — National Institute of Advanced Industrial Science and Technology

- Tomohiro Kudoh
- Hidemoto Nakada
- Atsuko Takefusa
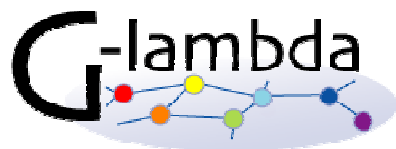- Yoshio Tanaka
- Fumihiro Okazaki
- Satoshi Sekiguchi

NiCT — National Institute of Information and Communications Technology

- Shuichi Okamoto
- Tomohiro Otani
- Yasunori Sameshima
- Atsushi Taniguchi
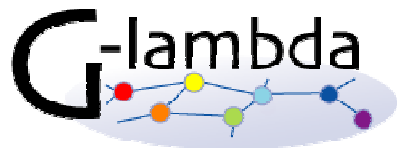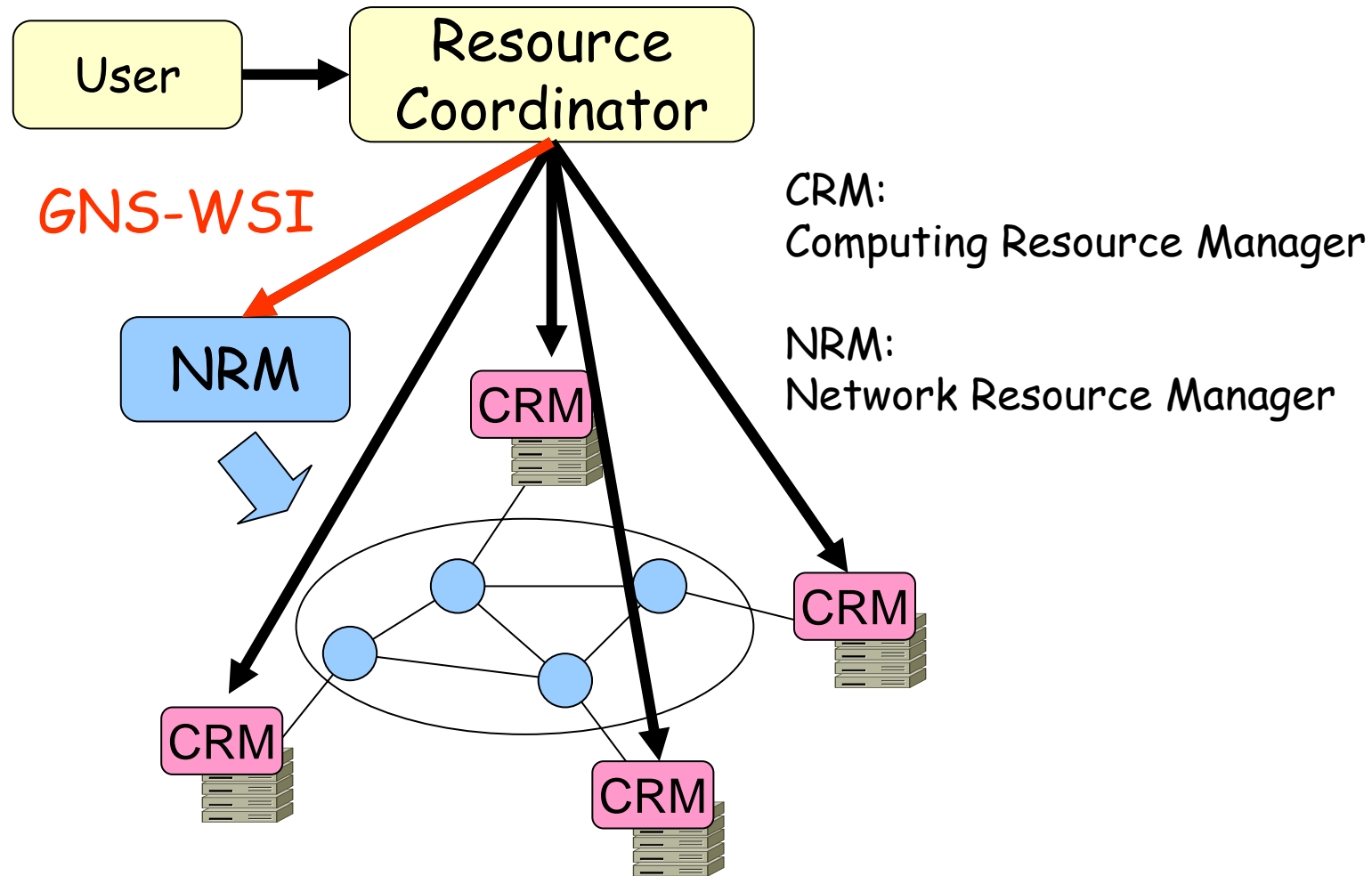
KDDI R&D LABS

- Michiaki Hayashi
- Takahiro Miyamoto
- Tomohiro Otani
- Hideaki Tanaka
- Masatoshi Suzuki

NTT

- A. Hirano
- Y. Sameshima
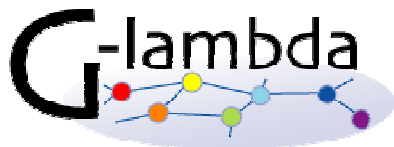- T. Ohara
- Y. Tsukishima
- A. Taniguchi
- M. Jinno
- Y. Takigawa

G-lambda system architecture

User → Resource Coordinator

GNS-WSI

NRM

CRM

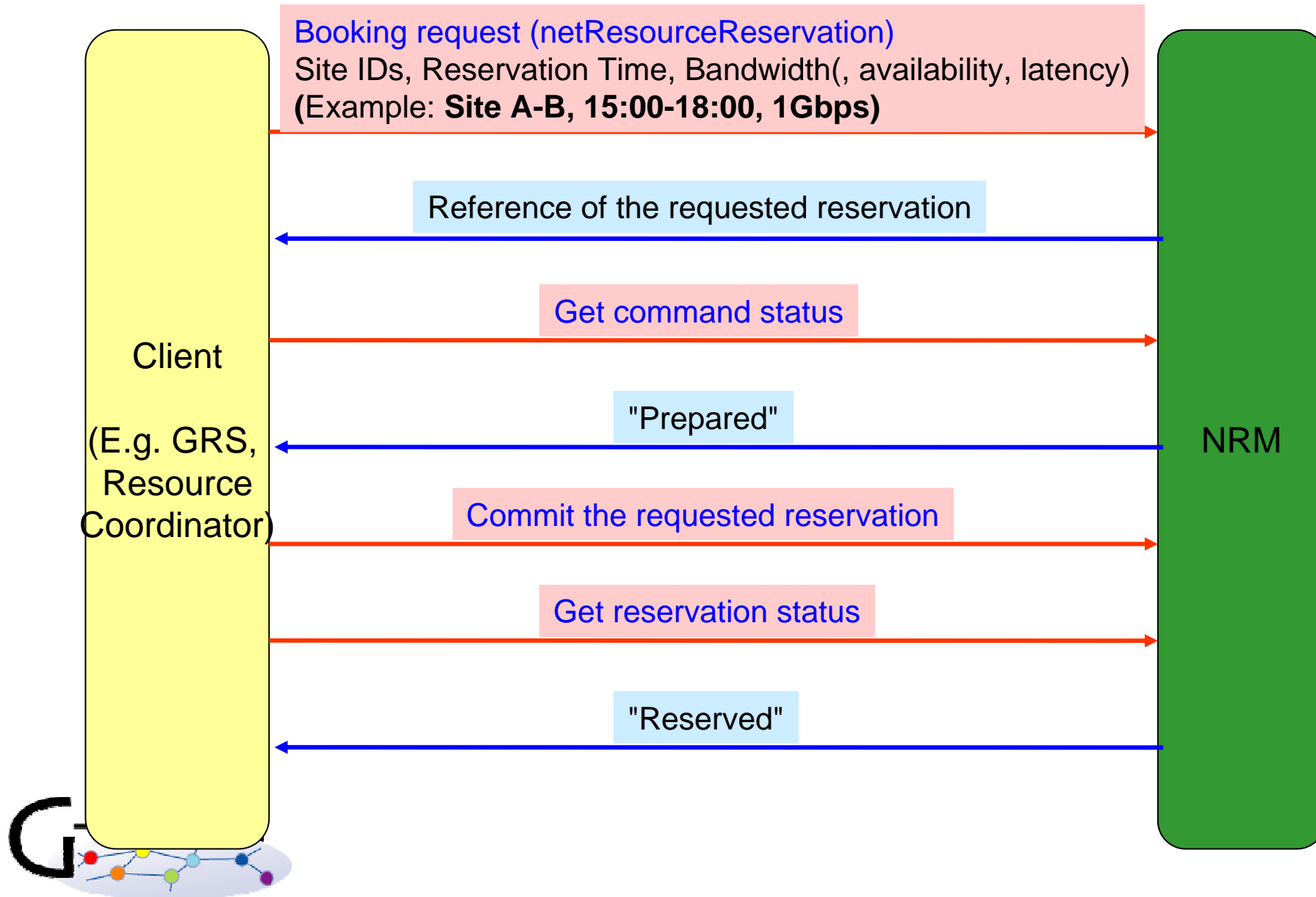CRM: Computing Resource Manager

NRM: Network Resource Manager

G-lambda

GNS-WSI (Grid Network Service / Web Services Interface)

- Grid Network Service-Web Services Interface
- Interface to realize advance reservation of bandwidth
- Based on the Web Services interface technology
- Can be used for inter-domain coordination
- Polling-based operations
  - Advance reservation of a path between end points
  - Modification of reservation (i.e. reservation time or duration)
  - Query of reservation status
  - Cancellation of reservation
- GNS-WSI2
  - WSRF(Web Services Resource Framework) based interface
    - GT4 (Globus Toolkit 4) Java WS Core http://www.globus.org/toolkit/
  - 2-phase commit

An example XML exchanged through GNS-WSI

Client

(E.g. GRS, Resource Coordinator)

NRM

Booking request (netResourceReservation)
Site IDs, Reservation Time, Bandwidth(, availability, latency)
(Example: **Site A-B, 15:00-18:00, 1Gbps)**

Reference of the requested reservation

Get command status

"Prepared"

Commit the requested reservation

Get reservation status

"Reserved"

Enlightened Computing

Enlightened Overview

Goals and Missions

- **Dynamic** and ***Adaptive*** on-demand and advanced reservation of end-to-end networking resources -

- Identification of functions and interactions between the control plane, management plane, and Grid middleware. Intelligent provisioning of lightpaths via middleware interactions with control plane protocols.

- To design and develop a Grid framework that provides applications/end-users to request computing and network resources in both a coordinated and reliable means. Create a virtualized set of resources dynamically.

- Determine how to abstract network resource information and how to distribute the network intelligence among the network control plane, management plane, and the Grid middleware.

- Feed near-real-time network state information to the Grid Resource Broker for optimized coordination and co-scheduling decisions

- Distributed Transaction Problem: ensure either the entire group of resources are successfully scheduled or none at all: *Highly-Available Robust Co-Scheduler* (HARC) open source, developed by CCT, LSU.

MCNC

G-lambda

en λIGHTened computing

R&D challenges

- The need to standardize the interfaces among Grid middleware and the network.
- Coordination and Co-scheduling of Network resources with other Grid resources (CPU, databases, sensors, instruments)
- Discovery, Monitoring, Adaptation system-level feedback control
- Extended L1/2 network services
  - On-demand vs. In-advance
  - Unicast, multicast, and anycast
- Control, management and middleware plane integration
  - GMPLS networking
  - Reconfiguration and re-optimization
  - Application controlled networking via the Grid middleware
- Testbed enabling dynamic service provisioning
  - GMPLS enabled PXC
  - E-NNI

The Enlightened Team

MCNC

- Yufeng Xin
- Steve Thorpe
- Bonnie Hurst
- Joel Dunn
- Gigi Karmous-Edwards
- Mark Johnson
- John Moore
- Carla Hunt
- Lina Battestilli
- Andrew Mabe

cct

- Ed Seidel
- Gabriele Allen
- Seung Jong Park
- Jon Maclaren
- Andrei Hutanu
- Lonnie Leger
- Dan Katz

STARLIGHT

- Joe Mambretti
- Alan Verlo

NC STATE UNIVERSITY

- Savera Tanwir
- Harry Perros
- Mladen Vouk

Calient NETWORKS

- Olivier Jerphagnon
- John Bowers

IBM
- Steven Hunter

renci renaissance computing institute

- Dan Reed
- Alan Blatecky
- Chris Heermann
- Ilia Baldin

MCNC

CISCO SYSTEMS EMPOWERING THE INTERNET GENERATION

- Javad Boroumand
- Russ Gyurek
- Wayne Clark
- Kevin McGrattan
- Peter Tompsu
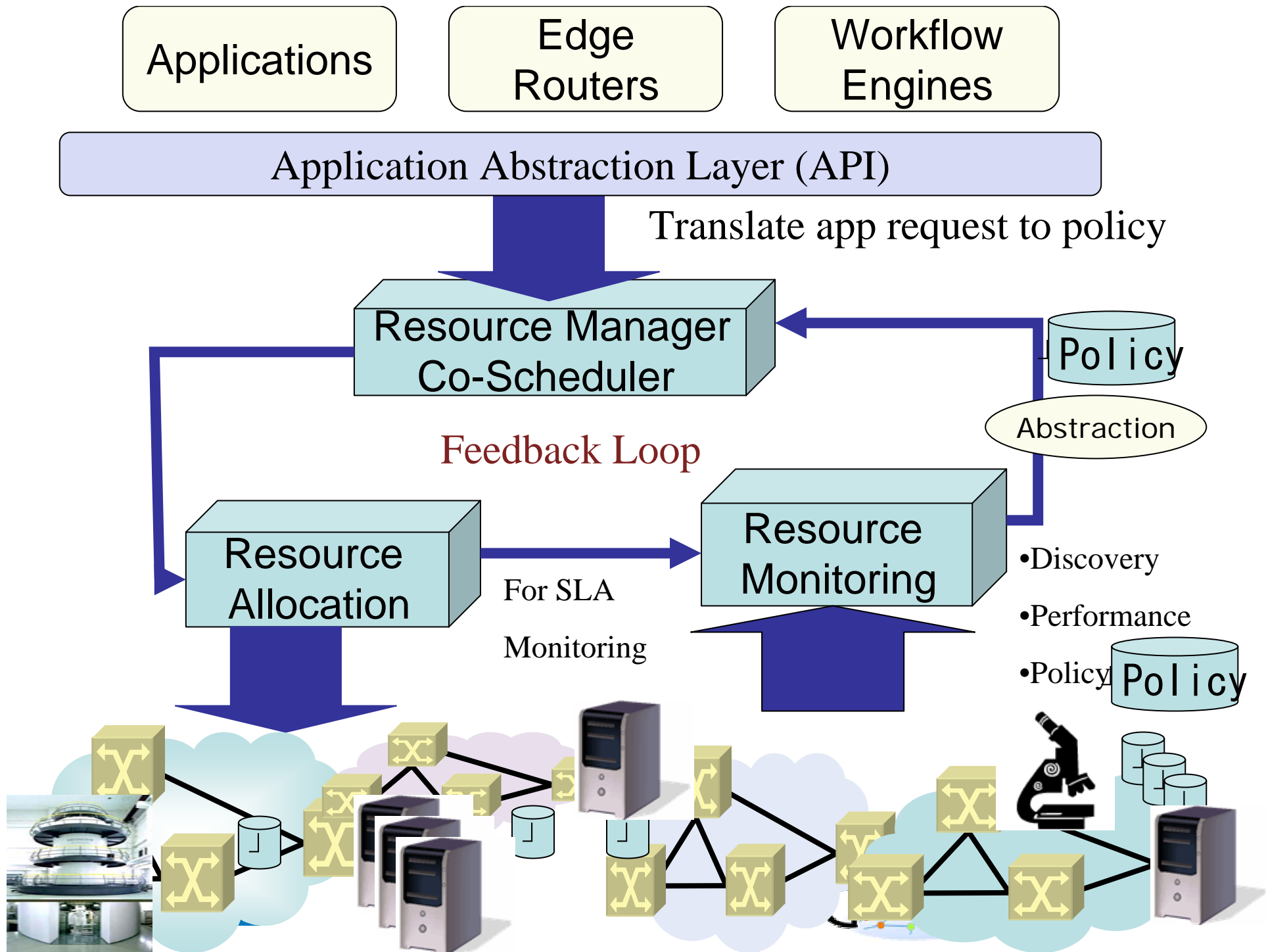
AT&T
- Rick Schlichting
- John Strand
- Matti Hiltunen

Naval Research Laboratory
- Yang Xia
- Xun Su

G-lambda

enλIGHTened computing

Applications

Edge Routers

Workflow Engines

Application Abstraction Layer (API)

Translate app request to policy

Resource Manager Co-Scheduler

Policy

Abstraction

Feedback Loop

Resource Allocation

For SLA Monitoring

Resource Monitoring

•Discovery

•Performance

•Policy

Policy

EnLIGHTened Resource Testbed

GMPLS Testbed

Control Channel

OSPF-TE

RSVP

Host 1

Eth SW

Eth SW

Host 2

PXC

OSPF-TE

RSVP-TE

LMP

PXC

MCNC

G-lambda
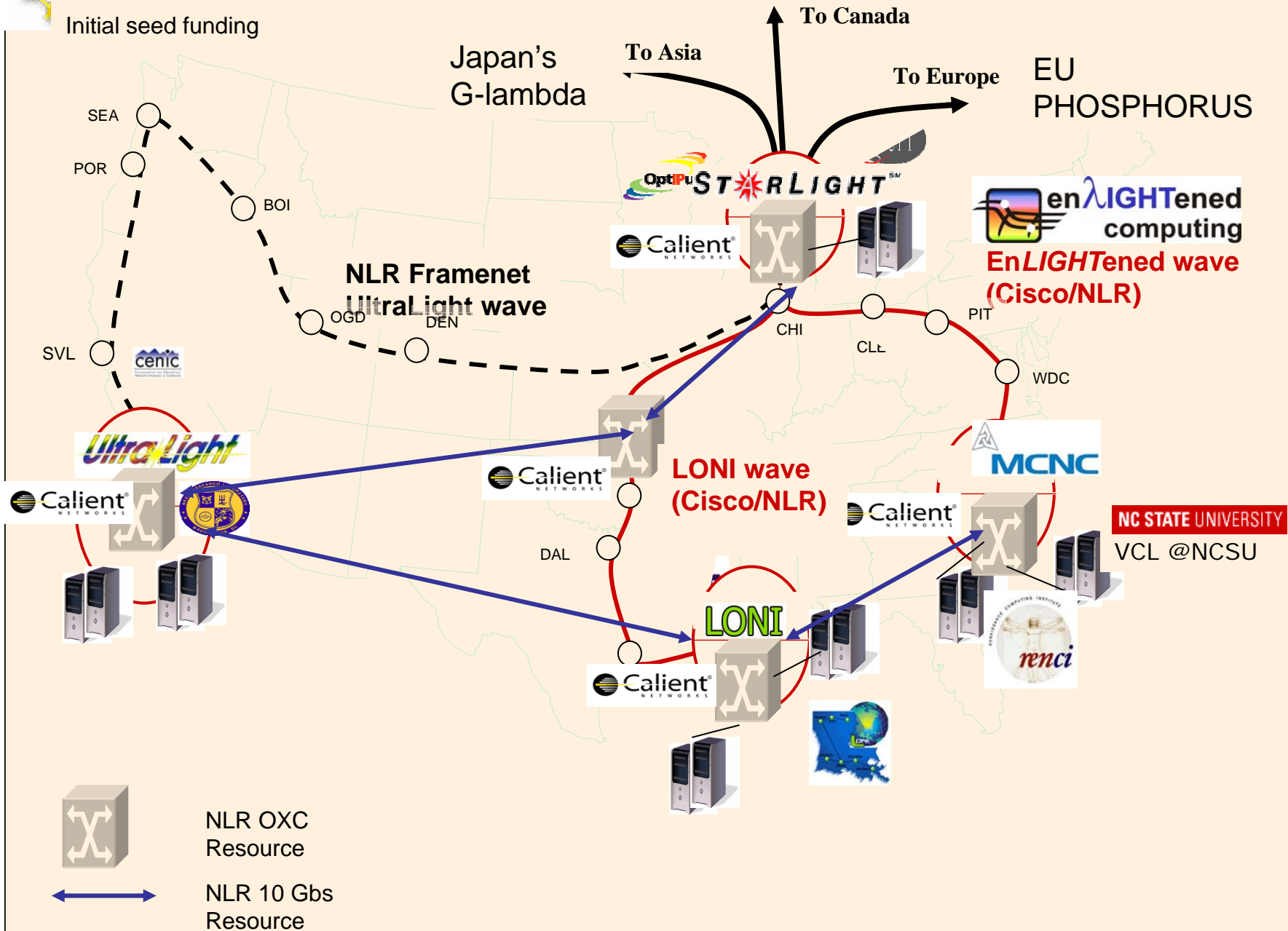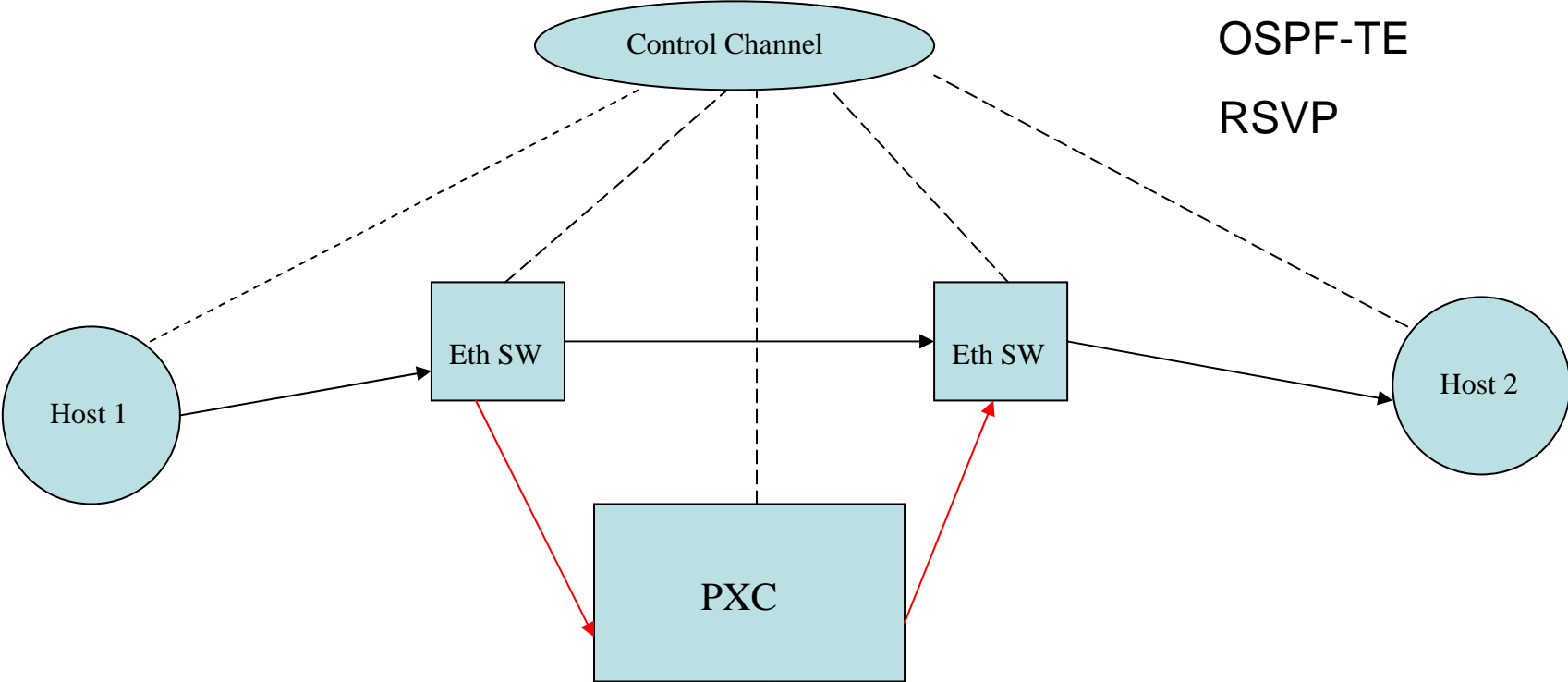
enλIGHTened computing
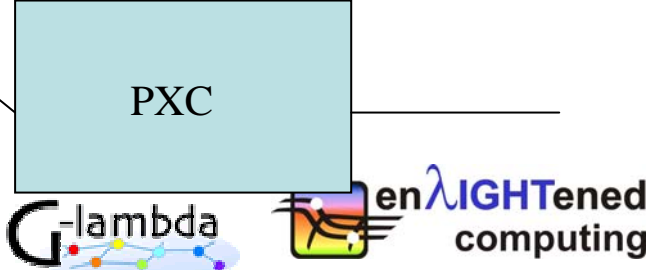
GMPLS E-NNI Demonstration

- Collaborative effort between NICT, NTT, KDDI Research and Enlightened Computing
  - Goal: to investigate potential for interdomain provisioning
- SC06 - demonstrated single-vendor interoperation between JGN II North (KDDI Research) and Enlightened
- December 2006 - tested three domain multi-vendor provisioning between JGN II South (NTT) and Enlightened, with JGN II North as transit domain.
- Configuration
  - Border nodes at Starlight, Otemachi-1(North) and Otemachi-2 (South)
  - Static external routes at border nodes point to other domain prefixes

Slide: John Moore

Architecture

HARC
Highly Available Robust Coallocator (LSU, Jon Maclaren)

- Extensible, open-sourced co-allocation system
- Can already reserve:
  - Time on supercomputers (advance reservation), and
  - Dedicated paths on GMPLS-based networks with simple topologies
- Uses 2-Phase Commit to atomically reserve multiple resources
- PaxosCommit used to provide reliable completion of bookings
- Was used to coordinate bookings across EnLIGHTened and G-lambda testbeds in largest demonstration of its kind to date.
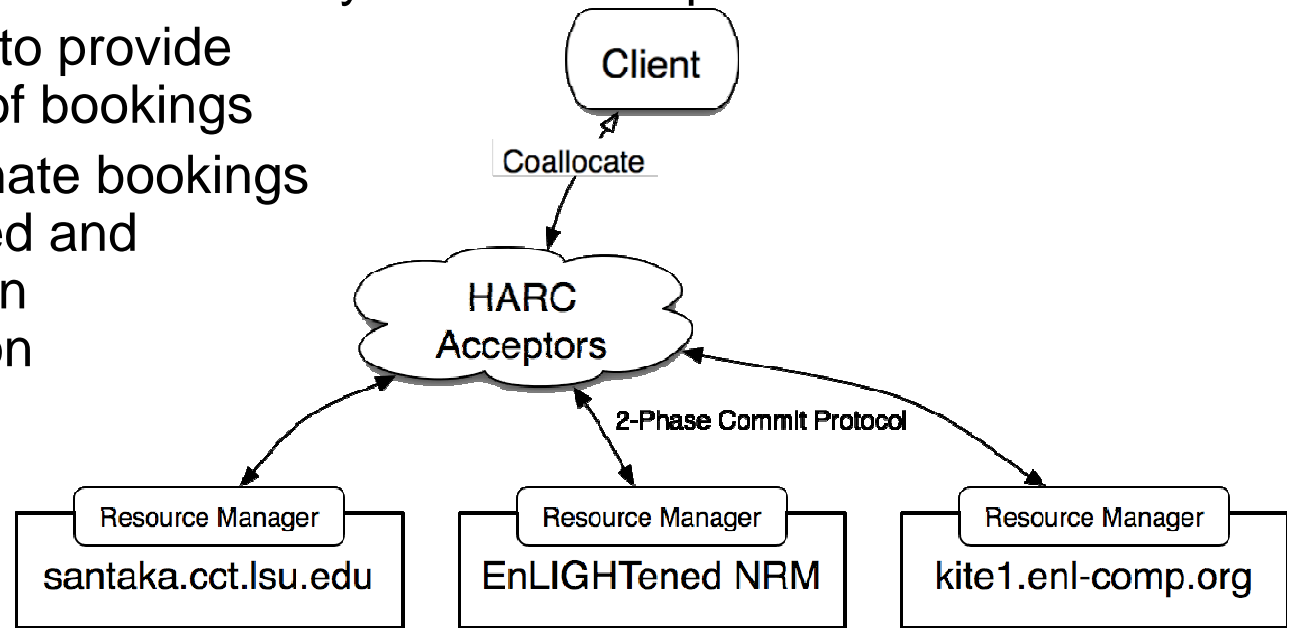
Slide: Jon Maclaren

MCNC

An HD-class example

LSU's Thomas Sterling HD-class on HPC

- 1.5 Gbs uncompressed stream each way
- Four Institutions participating, including (LSU, LaTech, University of Arkansas, Masatyk University (Czech Republic), MCNC/NCSU).
- Used Enlightened Middleware to schedule and set-up lightpaths for every Tuesday and Thursday session. (still need more work to make robust)

MCNC

G-lambda

en λIGHTened computing

NLR Partnership

NLR Partnership

- Establish a formal partnership based on a win-win model.
- Share knowledge, experience, and resources where appropriate.
- Create a working an NLR working group

  - We would like to have membership from several national and international projects
  - Work with NLR Research Council for input on requirements on the type of services NLR should provide to its constituents
  - Work with the international community on interconnecting NLR services with other NRENs
  - Develop a control plane and Grid middleware strategy with NLR to better serve the community
  - Utilize existing middleware and control plane software from Enlightened and other projects to help NLR users and Apps

Enlightened and G-lambda Collaboration

Japan's G-Lambda & EnLIGHTened collaboration
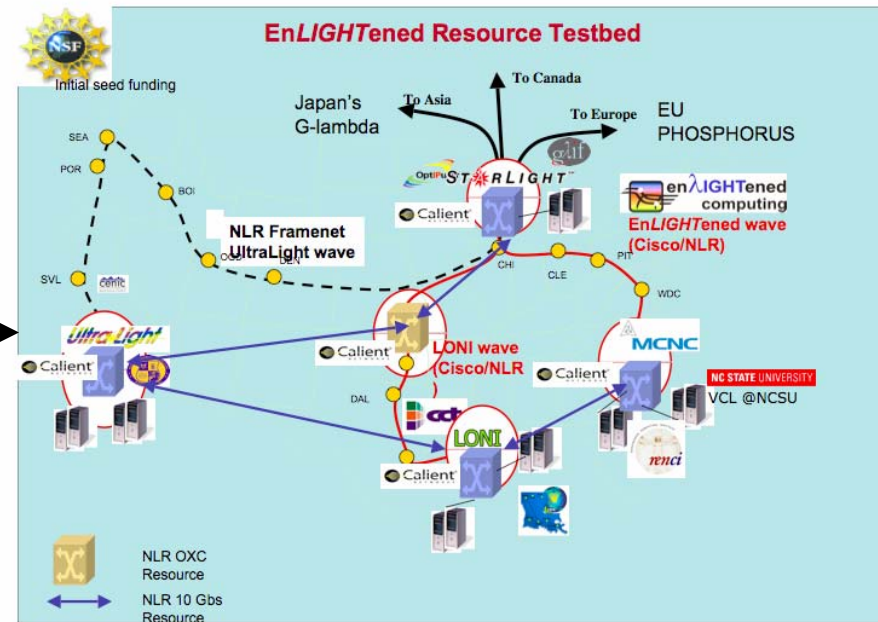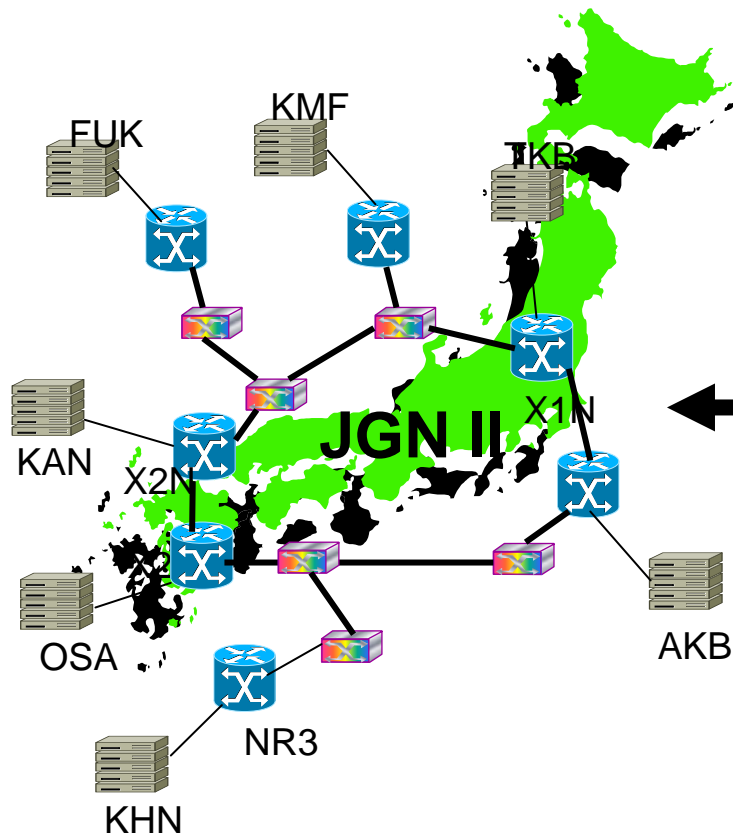
# Collaboration between G-lambda and Enlightened

What we achieved

- **Simultaneous in-advance reservation** of **bandwidth** between the US and Japan, and **computing resources** in the US and Japan
- **World's first** inter-domain coordination of resource mangers for in-advance reservation
  - Resource managers have different I/F and are independently developed
- "Automated" interoperability between network and computing resources in two countries' grid computing research testbeds
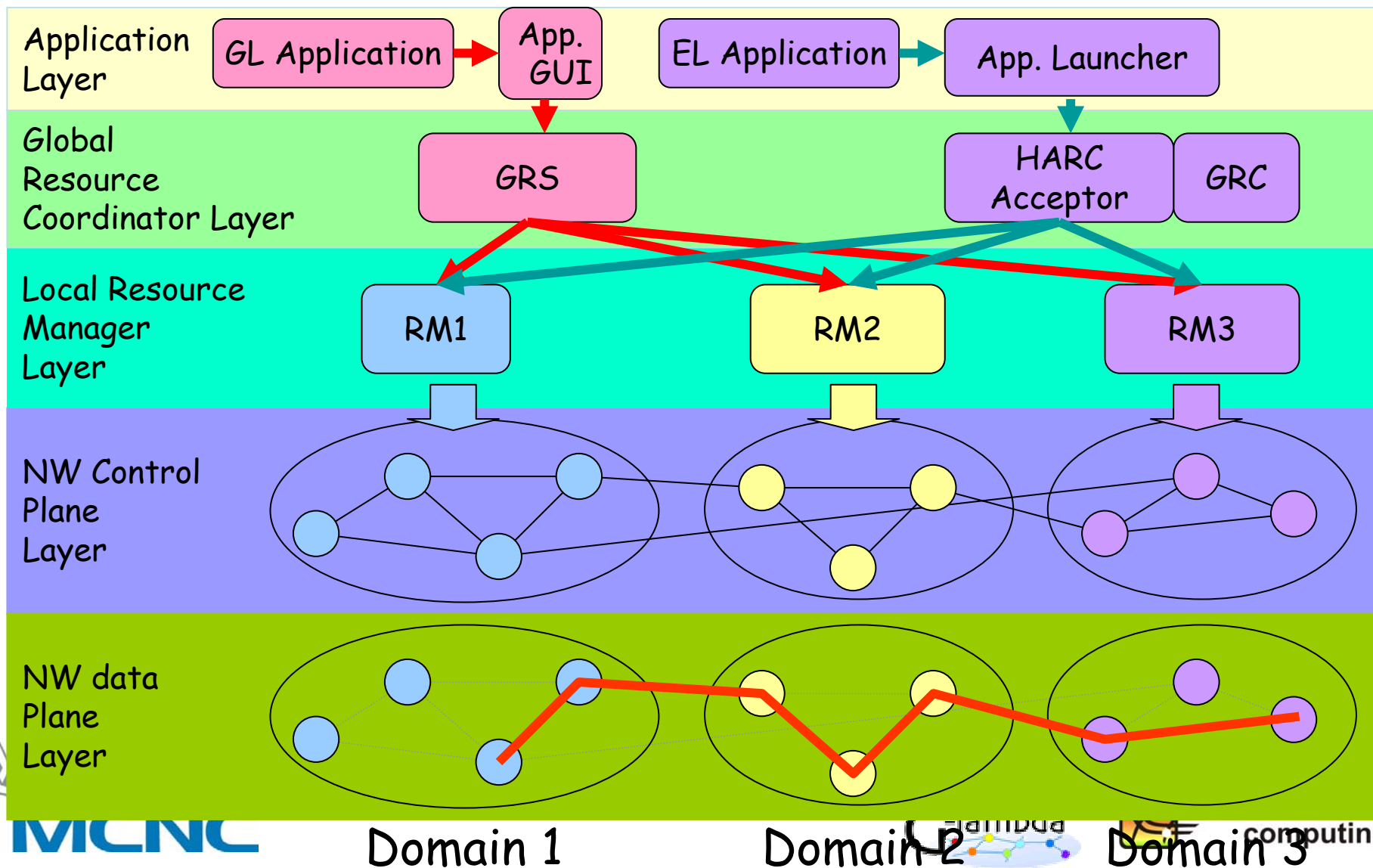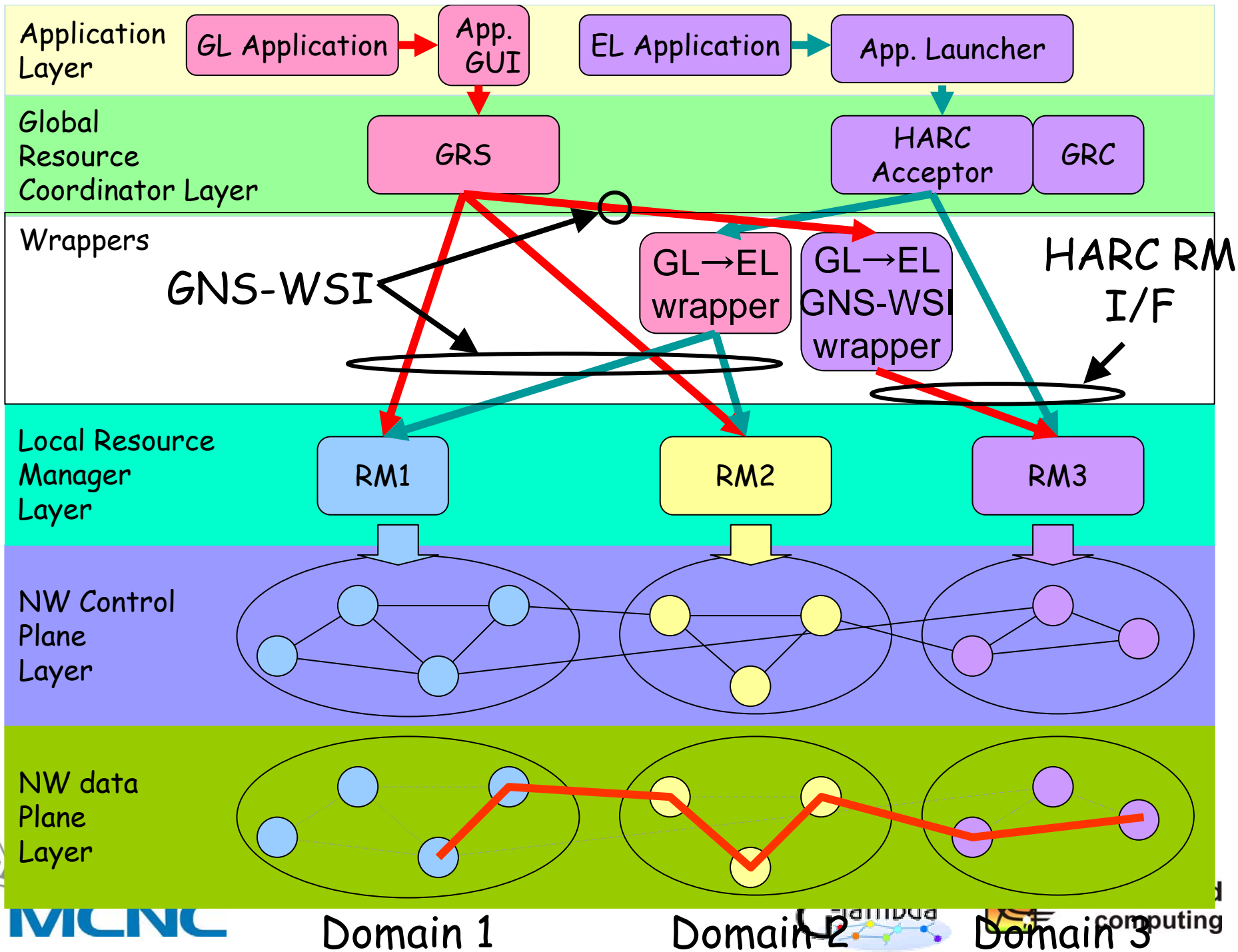
G-lambda/Enlightened middleware coordination diagram
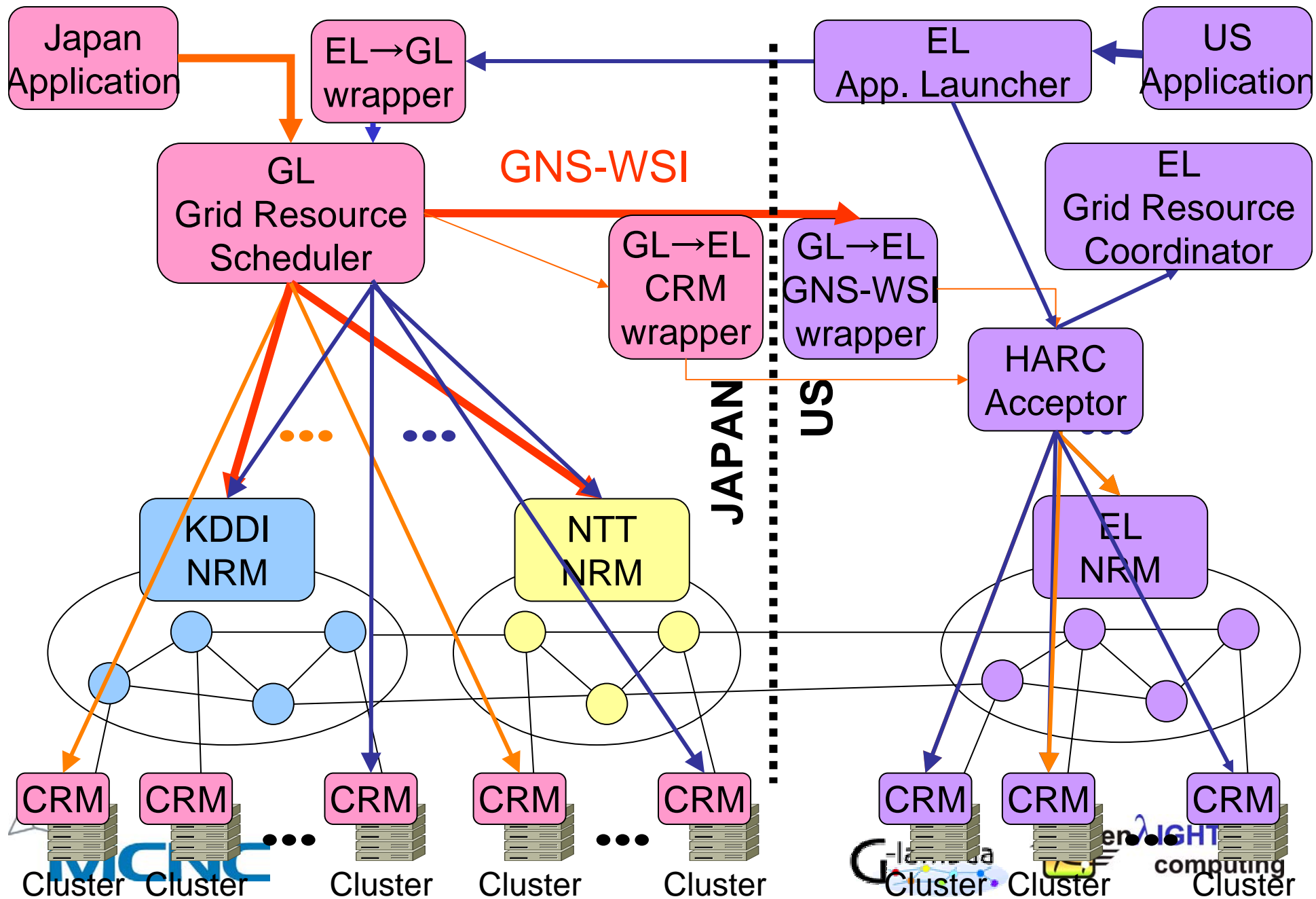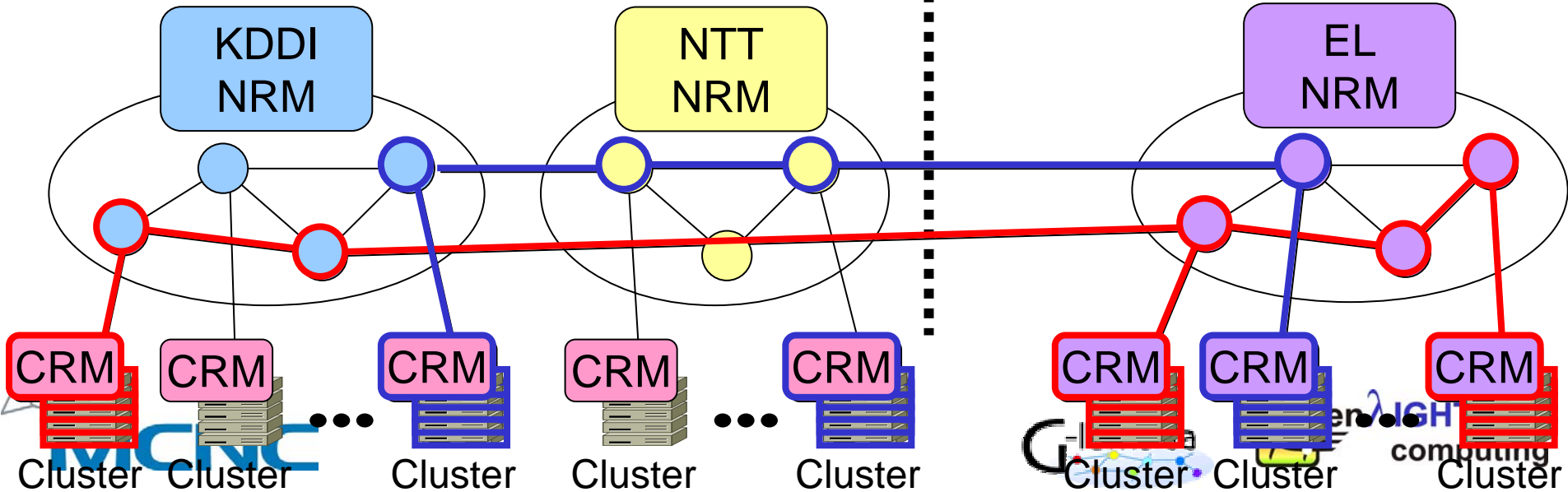
Resource map
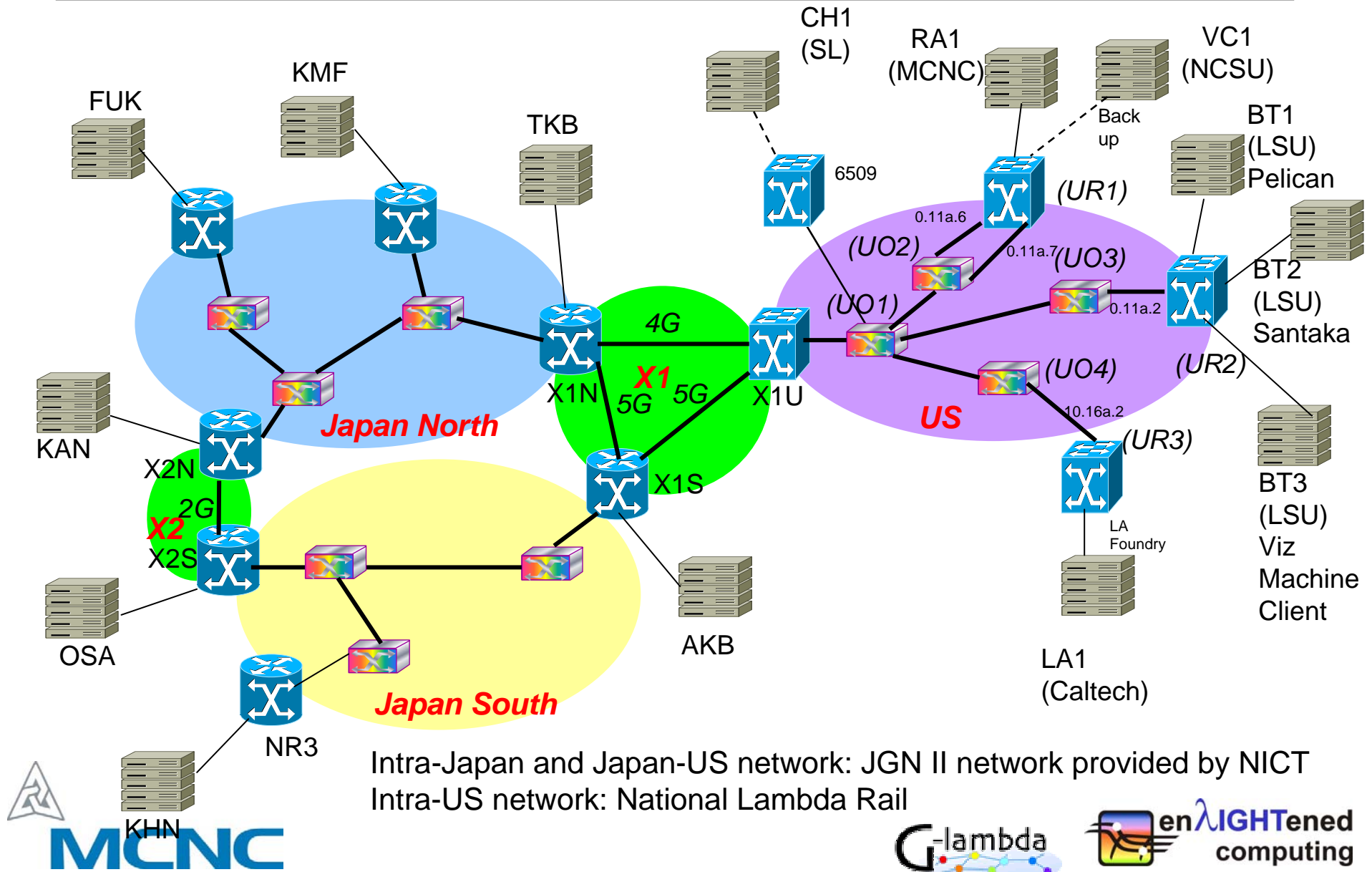
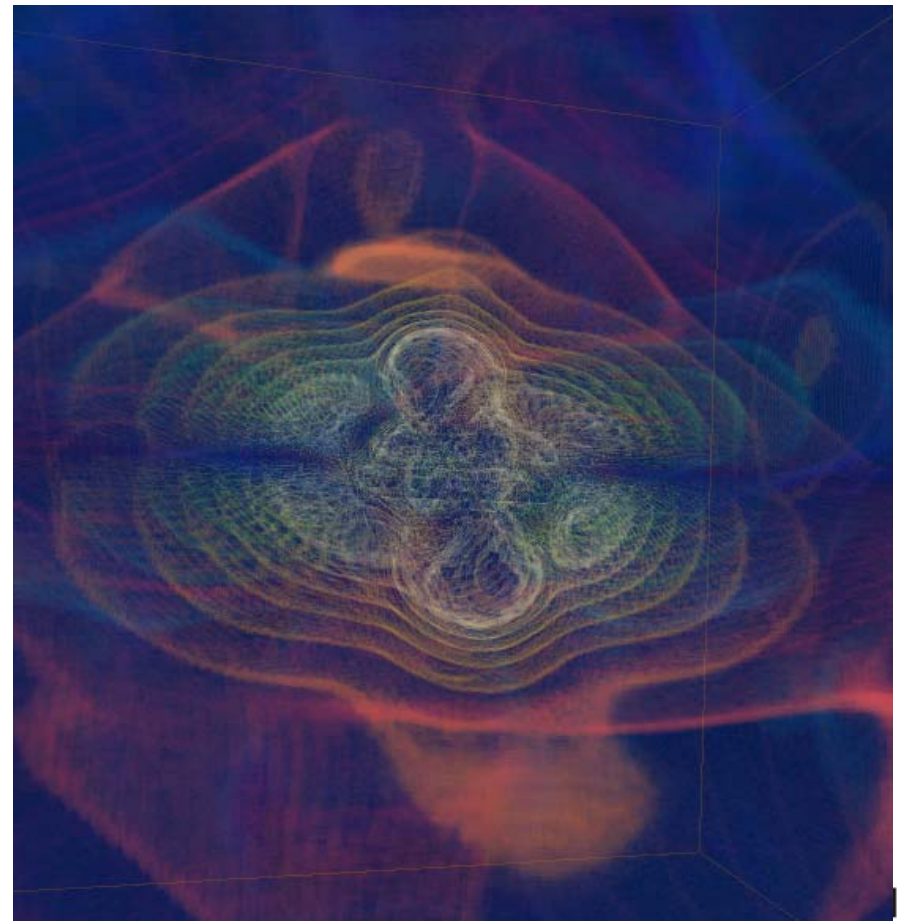Intra-Japan and Japan-US network: JGN II network provided by NICT
Intra-US network: National Lambda Rail

Enlightened: Visualization of remote data

- Data generated by remote simulation
  - Here : a black hole simulation
- Need to explore and visualize the dataset
- Enhanced Amira visualization system to take advantage of optical networks

MCNC

G-lambda   computing

Enlightened: Distributed data server

- Data available at multiple sites
- Distribution can be beneficial (parallelism, caching options, executing simple operations)
- A distributed data server (using the optical networks) can be faster than the local disk
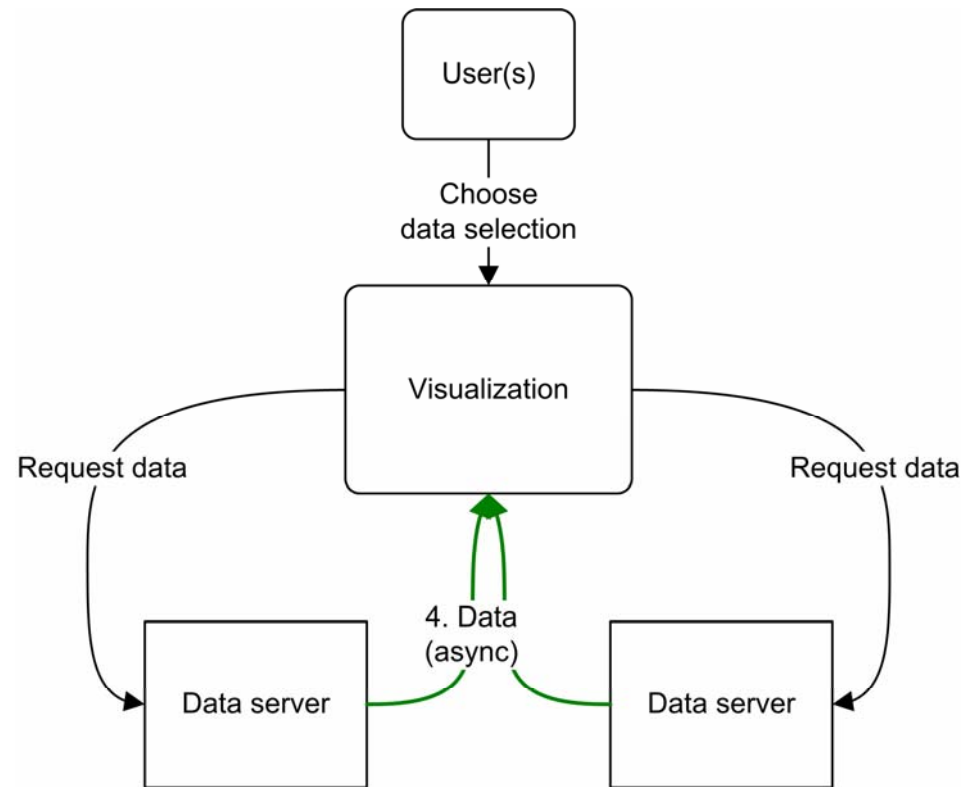
G-lambda: QM/MD simulation

- Surveying a chemical reaction path by Nudged Elastic Band method
  - calculating system configurations during the reaction in parallel

Start point

DEMO
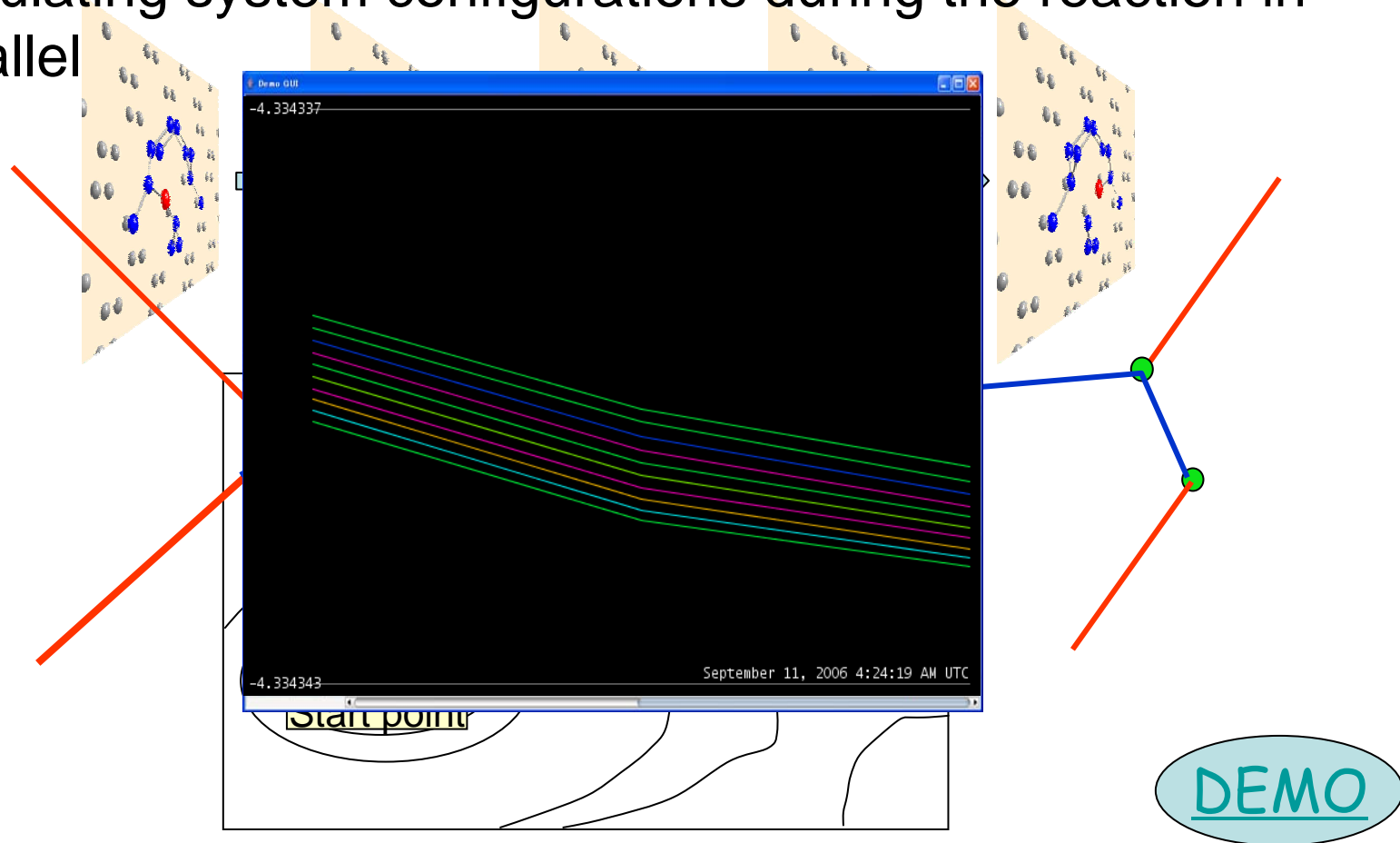
Three-Way collaboration

G-lambda, PHOSPHOUS, EnLIGHTened
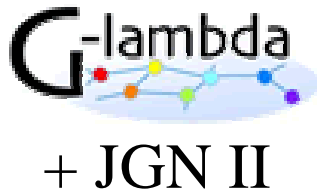
Three- Way Collaboration
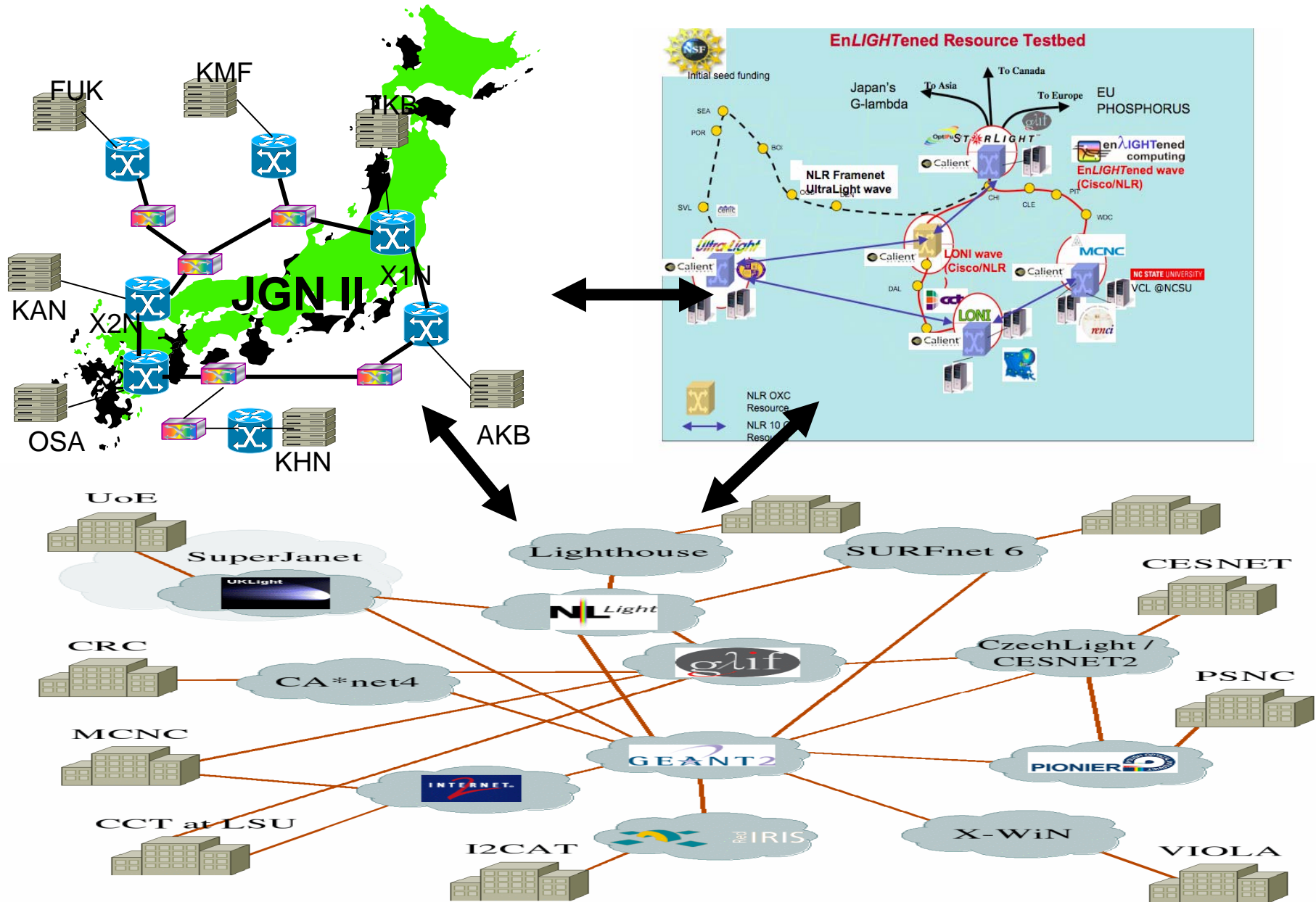
Collaboration of Three Continents meeting (C3C)

Who Are We and Why Are We Here?

Three Projects with similar Goals and mission

+ JGN II

- **We are collaborating because Grid is all about sharing of resources -and a Global Grid crosses all types of boundaries:**
  - **Administrative**
  - **Technological**
  - **Cultural**
  - **Funding Agencies**

Our challenges… agree on protocols and APIs

- **Currently we have several :**

- **Enlightened: XML + HTTP(S)**
- **G-lambda: XML + HTTP(S) + SOAP + WSRF**
- **Viola: XML + HTTP(S) + SOAP + WSRF**
  + WS-Agreement

  - How do we work towards an API for all of us to agree to implement to?
  - Each team will have different implementations of middleware
  - Each team will have different control for heterogeneous network technologies

Work Plan and Outcomes

- Work Plan and start of schedule:
  - Focus teams, middleware, testbeds, etc.
  - Conf call Schedule
  - Milestones
  - Next opportunity for F2F meeting
- Outcomes of this effort
  - Development of Two APIs agreed to
  - Integrated infrastructure
  - Available resources and middleware for advanced apps
  - WIKI site for C3C
  - Shared documents for GLIF and OGF, maybe IETF

GLIF

GLIF Control Plane and Grid Middleware Integration wg

1. Work with GLIF Tech group top establish what are GLIF resources (GOLEs) - done
2. Defined Network Elements in RDF - done
3. Software that reads RDF description - done
4. Need to write to Google MAP APIs to draw resources on a global bases - done

Need to work on

- Mechanisms to publish resource information
- APIs for requesting resources from NRMs (see next slide)
- So far the GLIF community has been about production resources
- Now introduce research testbeds , ie. Enlightened with policy
- Policy needs to be defined in human language then in machine

Thanks to Jereon Van Der Ham